A Neural Network Model to Predict Iron-Deficiency Anemia in Hemodialyzed Patients.

P. Baiardi, V. Piazza, G. Montagna, G. Franchi, M. Cortesi, M.C. Mazzoleni
Medical Informatics Unit, Division of Nephrology and Dialysis,
Salvatore Maugeri Foundation, Medical Center of Pavia, Italy.

Background. Patients with chronic renal failure treated by maintenance hemodialysis frequently suffer from iron-deficiency anemia (IDA). Even if serum ferritin (SF) levels are guidelines to iron therapy in these patients, other blood parameters should be considered to correctly evaluate their iron status. Indeed some patients, when given iron therapy, do not improve anemia in spite of low SF levels. The problem of diagnosing IDA in these patients may then be addressed to a multivariate pattern classification problem, and analytical methods may be a useful support to predict response to iron therapy.

We developed a neural network (NN) model to classify hemodialyzed patients as suffering or not from IDA, based on the assumptions of a multivariate nature of the problem, a not well defined behaviour of the involved variables and the presence of a non linear component in the problem itself. Method's performance was tested by comparison of results obtained from linear discriminant analysis (DA) and logistic regression (LR), methods with predetermined theoretical setting out, opposite to NN, not based on such an assumption.

Methods. We studied uremic patients on chronic hemodialysis at our Center in the last 5 years. After having checked medical eligibility criteria, we identified a subset of 52 patients from the total population, who, given iron therapy (more than once for some patients), produced 70 courses of therapy (cases). At least one year separated two consecutive courses, so cases were considered independent. Each case was classified as Not Responsive (NR, not suffering from IDA, identified by 0 in the model, 27 cases) and Responsive (R, suffering from IDA, identified by 1 in the model, 43 cases) to the therapy, according to a post-therapy variation in hemoglobin (Hb) levels. Hb values stable or improved of less than 1g/dl for NR, and improved of at least 2 g/dl for R. Predictor variables to IDA, i.e. input variables for the model, identified from clinical experience and literature, were: mean corpuscular volume, mean corpuscular Hb, mean corpuscular Hb concentration, serum iron, serum transferrin, transferrin saturation and SF.

A multilayer perceptron architecture of network with 1 hidden layer, logistic activation functions, linear net input to units and least squares objective function was chosen. Training technique was set to Levenber-Marquard algorithm, according to the small number of weights of the objective function. The optimal number of units in the hidden layer was identified according to Schwartz Bayesian Criterion of badness of fit. The final model was a 7-2-1 fully-connected architecture of network. Classification threshold for predicted values was arbitrary set to 0.5. Same criteria was used for LR results. The performance of the NN was evaluated by means of sensitivity (SE) and specificity (SP) in re-substitution (RE, model developed and tested on the entire data set) and cross-validation (CRV, data set randomly splitted into 3 developmental and validation data sets) experiments. Comparison with the traditional techniques was performed for both the RE and CRV experiments.

Results and conclusions. Both NN and traditional models well performed (SE>0.85), confirming that the problem is classifiable. In RE experiments NN correctly classify all the NR cases and 42/43 R cases (overall accuracy 0.99); in CRV experiments performances in general decreased (average SE=0.81, average SP=0.93), maintaining however high values of average accuracy (0.85). When compared with traditional approaches NN showed the highest values of SP in both RE (DA=0.92, LR=0.89) and CRV (DA=0.88, LR=0.85) testing, while the higher SE of RE experiments was not maintained in CRV. Reducing the false negative rate prevents the physician from failing to give iron replacement therapy when necessary. As a conclusion, we can state that a simultaneous taking into account of iron status parameters associated to a quantitative analytical method may ameliorate the correct classification of IDA patients. Our next goal is to achieve an ordinal classification of cases in different levels (low=1, medium=2, high=3) of IDA. An initial study on a data set of 58 patients for a total of 98 cases showed a fairly good performance of the NN and DA models, but not of the LR one. NN predicted extreme responses (1,3) with the same error rates of the binary model, while DA decreased performance (17/27 for NR and 32/43 for R). For the intermediate response DA acted better than NN. A more precise differentiation of levels of IDA, as suggested by more complex NN model tested, cannot be considered at the moment due to the esiguity of at disposal real cases.

1091—8280/97/$5.00 © 1997 AMIA, Inc.